BMC Immunology

Research article

Immunological hotspots analyzed by docking simulations: evidence for a general mechanism in pemphigus vulgaris pathology and transformation

Joo Chuan Tong*[1] and Animesh A Sinha[2]

Address: [1]Data Mining Department, Institute for Infocomm Research, 21 Heng Mui Keng Terrace, 119613, Singapore and [2]Center for Investigative Dermatology, Division of Dermatology and Cutaneous Sciences, College of Human Medicine, Michigan State University, 4120 Biomedical and Physical Sciences Building, East Lansing, MI 48824, USA

Email: Joo Chuan Tong* - jctong@i2r.a-star.edu.sg; Animesh A Sinha - asinha@msu.edu

* Corresponding author

Abstract

Background: Pemphigus vulgaris (PV) is an acquired autoimmune blistering disorder in which greater than 80% of active patients produce autoantibodies to the desmosomal protein desmogelin 3 (Dsg3). As the disease progresses, 40–50% of patients may also develop reactivity to a second component of the desmosomal complex, desmogelin 1 (Dsg1). T cells are clearly required for the production of autoantibodies in PV. However, few T-cell specificities within Dsg3 or Dsg1 have been reported to date, and the precise role of T-cells in disease pathogenesis and evolution remains poorly understood. In particular, no studies have addressed the immunological mechanisms that underlie the observed clinical heterogeneity in pemphigus. We report here a structure-based technique for the screening of DRB1*0402-specific immunological (T-cell epitope) hotspots in both Dsg3 and Dsg1 glycoproteins.

Results: High predictivity was obtained for DRB1*0402 (r^2 = 0.90, s = 1.20 kJ/mol, q^2 = 0.82, s_{press} = 1.61 kJ/mol) predictive model, compared to experimental data. *In silico* mapping of the T-cell epitope repertoires in Dsg3 and Dsg1 glycoproteins revealed that the potential immunological hotspots of both target autoantigens are highly conserved, despite limited sequence identity (54% identical, 72% similar). A similar number of well-conserved (18%) high-affinity binders were predicted to exist within both Dsg3 and Dsg1, with analogous distribution of binding registers.

Conclusion: This study provides interesting new insights into the possible mechanism for PV disease progression. Our data suggests that the potential T-cell epitope repertoires encoded in Dsg1 and Dsg3 is substantially overlapping, and it may be possible to apply a common, antigen-specific therapeutic strategy with efficacy across distinct clinical phases of disease.

Background

Pemphigus vulgaris (PV) is characterized by the loss of normal epithelial cell-to-cell adhesion leading to blistering which may involve the mucous membranes, non-

mucosal cutaneous surfaces, or both [1]. Pemphigus autoantibodies (autoAb) are mainly directed against the desmosomal glycoproteins desmoglein 3 (Dsg3) and des-

moglein 1 (Dsg1), members of the cadherin superfamily of cell adhesion molecules [2].

Clinical evolution of disease expression is common in PV [3,4]. In early disease, a majority of PV patients develop autoantibodies to Dsg3 coincident with mucosal blisters. In later stages, significant proportions of patients develop additional lesions on non-mucosal cutaneous sites and exhibit non-cross-reactive immunity to both Dsg3 and Dsg1 [5].

Two immunologic phenomenon termed "antigen mimicry" [5] and "epitope spreading" [5-8] have been proposed as possible pathogenic mechanisms responsible for the shift in autoreactive lymphocyte (T- or B-cell) profile from Dsg3+/Dsg1- to Dsg3+/Dsg1+. Antigen mimicry can be defined as the generation of lymphocyte (T- or B-cell) reactivity towards a protein due to its close structural similarity to unique exogenous antigens, or new determinants that have been generated endogenously [5]. Epitope spreading in the context of autoimmunity refers to the development of epitope-specific immune responses that are distinct from and non-cross-reactive with disease-inducing epitopes on the same (or different) protein secondary to the release of such a self-protein during an autoimmune response [8-10].

A close relationship between antigen mimicry and epitope spreading exists, with epitope spreading usually occurring after an initial episode of antigen mimicry [5]. Exogenous and endogenous antigens that may trigger cross-reactivity with self-proteins have not yet been defined in pemphigus [5]. While the modulation of autoantibody reactivities in the transformation of one disease subform into another has been actively explored [3-7], the role of T-cells underlying the evolution of autoreactive processes and epitope spreading remains poorly understood. To date, limited studies on T-cell specificities within PV have been reported [11-20]. The reported HLA associations with disease may serve to provide the genetic link that drives the evolving autoimmune responses in pemphigus. PV is known to be strongly associated with the HLA-DR allele DRB1*0402 [21-26]; it is present in more than 90% of Ashkenazi patients [27]. The DRB1*0402 allele is also common in other ethic backgrounds, including patients from France [28], Italy [29], Spain [30], Argentina [31] and Iran [32].

We have previously investigated the docking potentials of Dsg3 peptides to DRB1*0402 using a hybrid approach that integrates the strength of Monte Carlo simulations and homology modeling [33-37]. Consistent with experimental evidence [11], computational simulations reveal that a potentially large number of T-cell epitopes may be relevant in the pathogenesis of PV [33]. In the current study, we have extended our analysis to the Dsg1 glycoprotein and applied a new scoring scheme for identification of immunological (T-cell epitope) hotspots within both Dsg3 and Dsg1 self-antigens. *In silico* mapping of the T-cell epitope repertoires within Dsg3 and Dsg1 suggests that similar peptides from both PV target antigens may be involved in disease progression and the evolution in autoreactive lymphocyte reactivity during the course of disease from one clinical subtype to another (mucosal PV to mucocutaneous PV).

Results and Discussion
Comparison of Dsg1 and Dsg3 Extracellular Domains
The Dsg3 extracellular domain (ECD) has an extensive surface area of 32133 Å ^{2}. This surface area is ~3% larger than the Dsg1 ECD atomic accessible surface (31093 Å ^{2}). The general folds of Dsg1 and Dsg3 ECDs are similar (Figures 1 and 2). In particular, the ECD1, ECD2 and ECD3 of Dsg1 and Dsg3 are well conserved with Cα r.m.s.d. of 1.03 Å, 1.09 Å and 1.94 Å, respectively. The main differences between Dsg1 and Dsg3 ECDs lie in ECD4 (Cα r.m.s.d. = 3.76 Å) and ECD5 (Cα r.m.s.d. = 6.95 Å). In ECD4, the most obvious difference between the structures involve the loop 411–420 of Dsg1 and loop 405–422 of Dsg3. The backbone of Gln410, Ala411 and Ile412 in Dsg1 are intercalated more deeply into the ECD4 interface than the corresponding residues in Dsg3. In ECD5, the loop 468–479 curled upward toward Dsg1, with differences of up to 2.18 Å between the corresponding positions in Dsg3 (loop 468–475). It appears that these regions may not contain high concentration of T-cell epitopes and do not contribute directly to the differences in T-cell specificities of the two PV target antigens (Figure 3). Nonetheless, these structural differences are solvent exposed and may interact with pemphigus autoAb.

Dsg1 and Dsg3 Share Common Immunological Hotspots
The locations of Dsg1 and Dsg3 immunological hotspots are highly conserved across the ECD, TM and intracellular domain (ICD), despite limited sequence identity (54% identical, 72% similar). There are substantial overlaps between predicted and known immunological (T-cell epitope) hotspots (Dsg3 145–192, Dsg3 240–303, Dsg3 570–614) at the threshold -123 kJ/mol (Figure 2). At this threshold, the number of predicted Dsg3 hotspots are ten (residues 21–88, 125–147, 173–221, 245–299, 330–391, 435–456, 495–522, 584–640, 772–797, 830–859). Among these, one spans the signal peptide (SP), propeptide (PP) and the ECD; six are present in the ECD; one extends across the ECD, TM and ICD; while two hotspots are predicted to exist in the ICD (Figure 2). Similarly, nine hotspots were predicted for Dsg1 (residues 1–62, 82–151, 173–240, 310–372, 423–455, 470–498, 550–569, 656–688, 732–760), all of which overlap with those from Dsg3 (Figure 2). Collectively, these results suggest that the
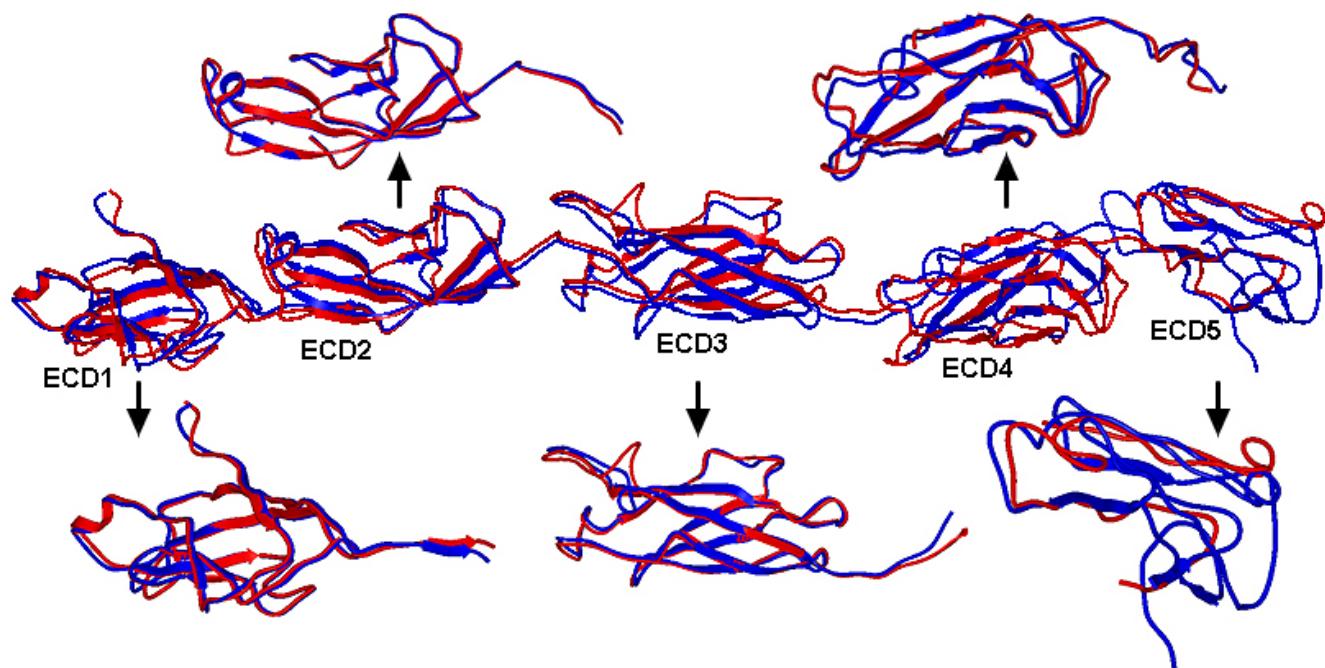
Figure 1
Superposition of the structures of Dsg1 and Dsg3 extracellular domains. Structural comparison of the Dsg1 (red) and Dsg3 (blue) extracellular domains. The r.m.s.d. of the entire ECD, ECD1, ECD2, ECD3, ECD4, and ECD5 are 5.56 Å, 1.03 Å, 1.09 Å, 1.94 Å, 3.76 Å, and 6.95 Å, respectively.

potential T-cell epitope repertoire encoded in Dsg1 and Dsg3 is substantially overlapping, and may help to explain the molecular basis underlying the observed intermolecular spreading from Dsg3 to Dsg1 targets during the course of PV.

Effects of Binding Registers in Peptide Selection
We investigated the effect of multiple contact regions or binding registers in Dsg3 peptides specific to DRB1*0402 [33]. Among 985 Dsg3 peptide sequences (including SP and PP derived sequences), 658 were predicted high-affinity binders with 77% containing two or more registers that can be docked into the binding groove of DRB1*0402. A similar number of high-affinity binders are predicted to exist within the Dsg1 glycoprotein, with analogous distribution of binding registers as illustrated in Figure 4. Of 1035 Dsg1 peptides, 665 were predicted high-affinity binders with 129 (~23%), 112 (~19%), 130 (~17%), 99 (~15%) and 43 (~6%) peptides possessing two, three, four, five and six registers, respectively.

Redundancy Profiles of Predicted Binding Peptides
The predicted Dsg1 high-affinity binding sequences were examined for their similarity with the Dsg3 proteome. Each Dsg1 (15 mer) sequence was used to probe the entire Dsg3 proteome for the highest identity Dsg3 (15 mer) peptide with the minimal number of substitutions. Figure 5 details the degree of conservation of Dsg1 predicted high-affinity binders with Dsg3 sequences. All predicted Dsg1 and Dsg3 (15 mer) binding peptides share at least four common residues along their primary sequences. It has been reported that PV and PF autoAbs can cross-react with Dsg1 and Dsg3 peptides with 75% identity [38]. Sequence alignment showed that 18% (or 122) of these peptides are highly conserved with at least 75% sequence identity. In this context, two peptides $Dsg1_{58-72}$ CREGEDNSKRNPIAK and $Dsg1_{59-73}$ REGEDNSKRNPI-AKI near the N-terminus of ECD1 appear to be of particular interest, since they are fully conserved within the Dsg3 proteome and may represent the most likely antigenic link between self-directed responses to two distinct autoantigens (Dsg3 and Dsg1). A subset of patients with PV has been shown to have T cell reactivity to both Dsg3 and Dsg1 [39]. However, defined Dsg1 epitopes have not yet been determined.

Conclusion
Although clinical evolution of disease expression is common in PV [3,4], our understanding of how T-cells are involved in the underlying autoreactive processes remains poor. Collectively, the results of this study provide interesting new insights into the possible mechanism of dis-
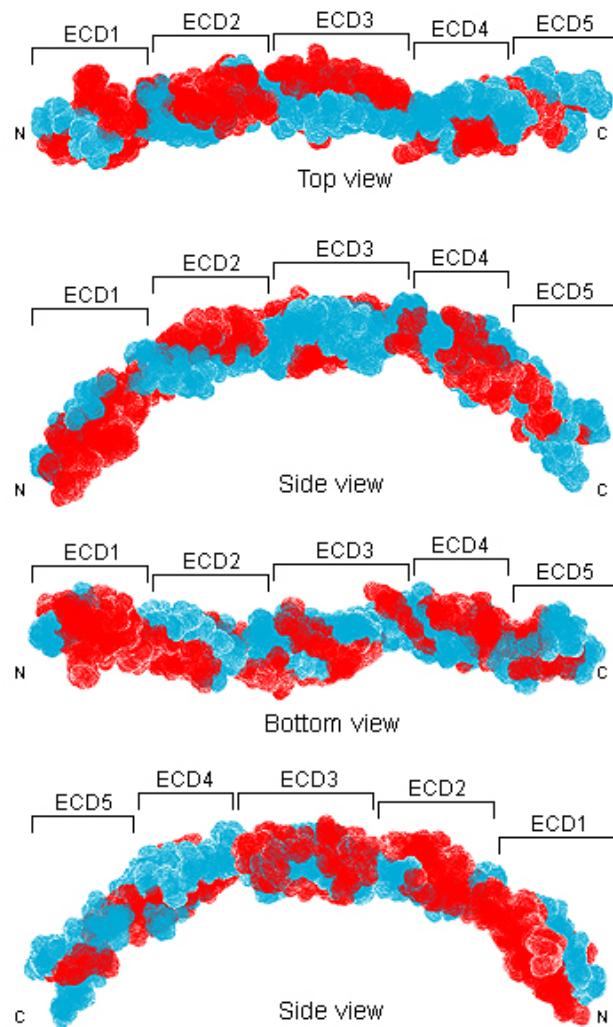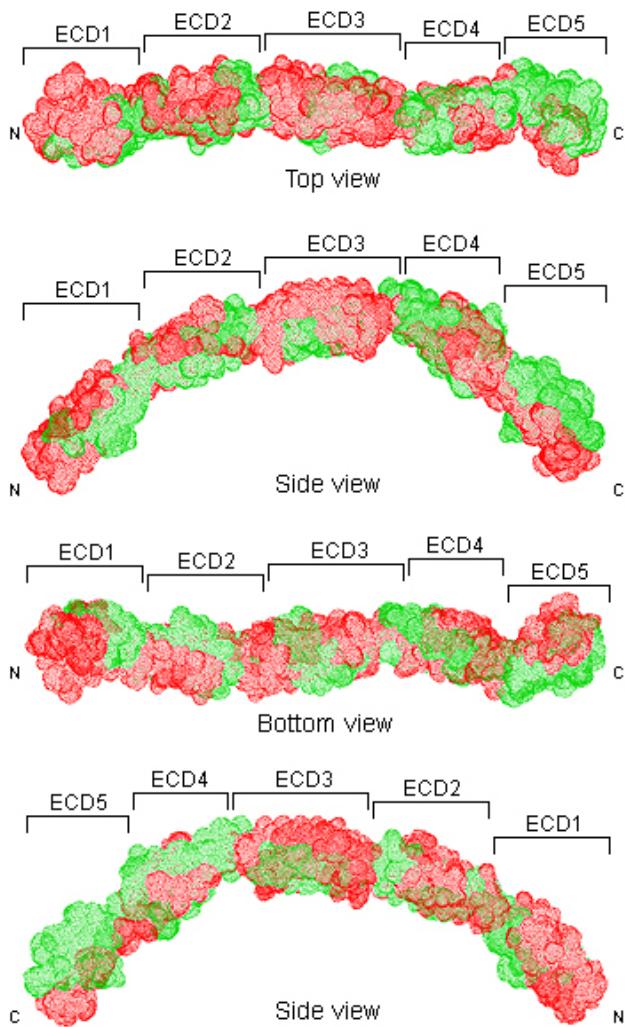
Figure 2
Molecular surface of Dsg1 and Dsg3 extracellular domains illustrating the positions of predicted DRB1*0402-specific immunological (T-cell epitope) hot spots. Locations of predicted immunological hot spots (colored red) for Dsg1 and Dsg3 extracellular domains are shown in surrounding views and numbered in accordance to their extracellular domains.

ease progression. Our data suggests that similar peptides from the two known PV autoantigenic targets may bind to DRB1*0402 and allow for intermolecular epitope spreading that lead to distinct morphological categories of PV – mucosal lesions only *vs.* mucocutaneous disease. Nonetheless, many other factors exist, such as IgG-activated intracellular signalling events [40], which may play a critical role in the complex disease mechanism and should also be explored. Recently, Lucchese and coworkers [38] discovered an immunodominant Dsg3 T-cell epitope Dsg3$_{49-60}$ REWVKFAKPCRE that is highly reactive with both PV and PF sera, with 75% identity to Dsg1$_{49-60}$

REWIKFAAACRE. In this context, our current approach of epitope mapping may prove useful in facilitating the systematic discovery of peptides that cross-react with both PV and PF autoreactive T-cells and serve as targets for potential therapeutic approaches that are efficient in both diseases. Further studies are necessary to determine the proportion of immunogenic epitopes that are capable of stimulating autoreactive lymphocytes from both pemphigus subtypes.

```
Dom    |         SP        |          PP         |          ECD1
Pos    1........10........20........30........40........50........60.......70
Dsg1   MDWSFFRVVAVLFIFLVVVEVNSEFRIQVRDYNTKNGTIKWHSIRRQKREWIKFAAACREGEDNSKRNPI
Dsg3   MMGLFPRTTGALAIFVVVILVHGELRIETKGQYDEEEMTHQQAKRRQKREVVKFAKPCREGEDNSKRNPI
       *    * *....* **:**: *:.*:**::.  :     :: *******:*** .**************

Dom                            ECD1
Pos    71.......80........90........100.......110.......120.......130.....140
Dsg1   AKIHSDCAANQQVTYRISGVGIDQPPYGIFVINQKTGEINITSIVDREVTPFFIIYCRALNSMGQDLERP
Dsg3   AKITSDYQATQKITYRISGVGIDQPPFGIFVVDKNTGDINITAIVDREETPSFLITCRALNAQGLDVEKP
       *** ** *.*.;*:************;****:;;:**;****:***** ** *;* *****: * *;*:*

Dom    141......150.......160.......170.......180.......190.......200.....210
Pos                     |                            ECD2
Dsg1   LELRVRVLDINDNPPVFSMATFAGQIEENSNANTLVMILNATDADEPNNLNSKIAFKIIRQEPSDSPMFI
Dsg3   LILTVKILDINDNPPVFSQQIFMGEIEENSASNSLVMILNATDADEPNHLNSKIAFKIVSQEPAGTPMFL
       * * *;:************* ** *;***** :*:*********;********* ***:;****:

Dom    211......220.......230.......240.......250.......260.......270.....280
Pos                       ECD2                           |       ECD3
Dsg1   INRNTGEIRTMNNFLDREQYGQYALAVRGSDRDGGADGMSAECECNIKILDVNDNIPYMEQSSYTIEIQE
Dsg3   LSRNTGEVRTLTNSLDREQASSYRLVVSGADKDG--EGLSTQCECNIKVKDVNDNFPMFRDSQYSARIER
       :.***;**:.* ***** .;* *.* *;*:** :*;*:;******: ***** :*;.;*.: .*;*

Dom                             ECD3
Pos    281......290.......300.......310.......320.......330.......340.....350
Dsg1   NTLNSNLLEIRVIDLDEEFSANUMAVIFFISGNEGNWFEIEMNERTNVGILKVVKPLDYEAMQSLQLSIG
Dsg3   NILSSELLRFQVTDLDEEYTNWLAVYFFTSGNEGNWFEIQTDPRTNEGILKVVKALDYEQLQSVKLSIA
       * *.**.::* ***** ;*** ** ***; *:** ** *** ************;**** :**:;***.

Dom    351......360.......370.......380.......390.......400.......410.....420
Pos                        ECD3                    |                 ECD4
Dsg1   VRNKAEFHHSIMSQYRLKASAISVTLNVIEGPVFRPGSKTYVVTGNMGSNDK-VGD-FVAT----DLDT
Dsg3   VKNKAEFHQSVISRYRVQSTPVTIQVINVREGIAFRPASKTFTVQKGISSK-KLV-DYILGTYQAIDEDT
       *:******;*:*:*;*:::::;:::  *;** ** .***.***;.* ..;*  * * * ::**   * **

Dom    421......430.......440.......450.......460.......470.......480.....490
Pos             ECD4                                              |
Dsg1   GRPSTTVRYVMGNNPADLLAVDSRTGKLTLNKVTKEQYNMLGGKYQGTILSIDDNLQRTCTGTININIQ
Dsg3   NKAASNVKYVMGRNDGGYLMIDSKTAEIKFVKNMNRDSTFIVNKTITAEVLAIDEYTGKTSTGTVYVRVP
       .:.::;.*;****.* .. * ;**;*.::.; :::.;:. ::. . .;*;**:  ;*.***: :.:

Dom                    ECD4         |                     ECD5
Pos    491......500.......510.......520.......530.......540.......550.....560
Dsg1   SFGN---------DDRTNTEPNTKITTNTGR----------------------------------QES
Dsg3   DFNDNCPTAVLEKDAVCSSSPSVVVSARTLNNRYTGPYTFALEDQPVKLPAVWSITTLNATSALLRAQEQ
       .*.:      *   .:.*.. :::.* .                                      **.

Dom                                  ECD5                          |      TM
Pos    561......570.......580.......590.......600.......610.......620.....630
Dsg1   TSSTNYDTSTTSTDSSQVYSSEP----------------------GNGADKDLLSDNVHFGPAGIGLLIM
Dsg3   IPPGVYHISLVLTDSQNNRCEMPRSLTLEVCQCDNRGICGTSYPTTSPGTRYGRPHSQRLGPAAIGLLLL
       .. *.*. . ***.: .. *             . *:: ... :;***.****::
```

```
Dom         TM      |                         ICD
Pos    631......640.......650.......660.......670.......680.......690.....700
Dsg1   GFLVLGLVPFLMICCDCGGAPRS--AAGFEPVPECSDGAIHSWAVEGPQPEPRDITTVIPQIPPDNANII
Dsg3   GLLLLLLAPLLLLTCDCGAGSTGGVTGGFIPVPDGSEGTIHQWGIEGAHPEDKEITNIC--VPPVTANGA
       *:*;* *.*;*:: ****... .  :.** ***: *:*:**.*.:**.:** ::**.:   :** .**

Dom                                   ICD
Pos    701......710.......720.......730.......740.......750.......760.....770
Dsg1   ECIDNSGVYTNEYG---------GREMQDLGGGERMTG--FELTEGVKT--------SGMPEICQEYSG
Dsg3   DFMESSEVCTNTYARGTAVEGTSGMEMTTKLGAATESGGAAGFATGTVSGAASGFGAATGVGICSSGQSG
       : ::.* * ** *.     * **  *.  :* :: *. :        :*:  .. **

Dom    771......780.......790.......800.......810.......820.......830.....840
Pos                              ICD
Dsg1   TLRR-----NSMRECREGGLNMNFMESYFCQKAYAYADEDEGRPSNDCLLIYDIEG---VGSPAGSVGCC
Dsg3   TMRTRHSTGGTNKDYADGATSWNFLDSYFSQKVFACAEEDDGQEANDCLLIYDNEGADATGSPVGSVGCC
       *:*     .: :: ;*.:.***;:***.***;* *;**;*; ;******** ** .***.******

Dom    841......850.......860.......870.......880.......890.......900.....910
Pos                               ICD
Dsg1   SFIGEDLDDSFLDTLGPKFKKLADISLGKESYPDLDPSWPPQSTEPVCLPQETEPVVSGHPPISPHFGTT
Dsg3   SFIADDLDDSFLDSLGPKFKKLAEISLGVD---------------------GEGKEVQ---PPSK-----
       ***.;******;*********;**** :                       *. * *

Dom    911......920.......930.......940.......950.......960.......970.....980
Pos                               ICD
Dsg1   TVISESTYPSGPGVLHPKPILDPLGYGNVTVTESYTTSDTLKPSVHVHDNRPASNVVVTERVVGPISGADL
Dsg3   ----DSGYG-----------IESCGHPIEVQQTGFVKCQTLS-----G-SQGASALSASGSVQPAVS----
       :* *            ::. *:   . .:::.:***.       .; ** : .:  *  .:*

Dom    981......990.......1000......1010......1020......1030......1040....1050
Pos                               ICD
Dsg1   HGMLEMPDLRDGSNVIVTERVIAPSSSLPTSLTIHHPRESSNVVVTERVIQPTSGMIGSLSMHPELANAH
Dsg3   -----IPDPLQHGNYLVTETYSASGSLVQPSTAGFDPLLTQNVIVTERVICPISSVPGNLAGPTQLRGSH
         :**  : .* ;*** *...* : .* : ..* ;.**;****** * *;.: *.*: .:* .;*

Dom    1051....1060......1070......1080......1090......1100......1110....1120
Pos                               ICD
Dsg1   NVIVTERVVSGAGVTGISGTTGISGGIGSSGLVGTSMGAGSGALSGAGISGGGIGLSSLGGTASIGHMRS
Dsg3   TMLCTEDPCS------------------R-LI-------------------------------------
       .:; ** *                    *:

Dom                                   ICD
Pos    1121....1130......1140......1150
Dsg1   SSDHHFNQTIGSASPSTARSRITKYSTVQYSK
Dsg3   --------------------------------
```

Figure 3

Predicted DRB1*0402-specific immunological hot spots for Dsg1 and Dsg3 glycoproteins. Multiple sequence alignment of Dsg1 and Dsg3. Predicted hot spots (threshold = -123 kJ/mol) shaded in black, experimental Dsg3 hot spots highlighted red, and experimentally confirmed DRB1*0402-specific Dsg3 peptides highlighted blue. Dsg1 – Signal Peptide, SP: 1–23; Propeptide, PP: 24–49; Extracellular domain 1, ECD1: 50–158; ECD2: 159–270; ECD3: 271–385; ECD4: 386–497; ECD5: 498–548; Transmembrane, TM: 549–569; Intracellular domain, ICD: 570–1049. Dsg3 – SP: 1–23; PP: 24–49; ECD1: 50–158; ECD2: 159–268; ECD3: 269–383; ECD4: 386–499; ECD5: 500–615; TM: 616–640; ICD: 641–999.
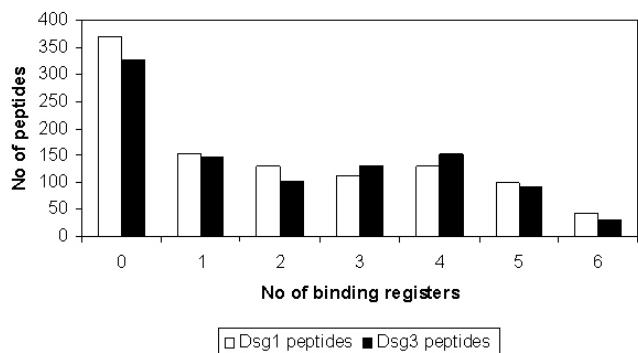
Figure 4
Number of predicted binding registers for Dsg1 and Dsg3 peptides bound to DRB1*0402. The frequency of Dsg1 (colored white) and Dsg3 (colored black) peptides docked to DRB1*0402 shown as a function of the number of predicted binding registers.
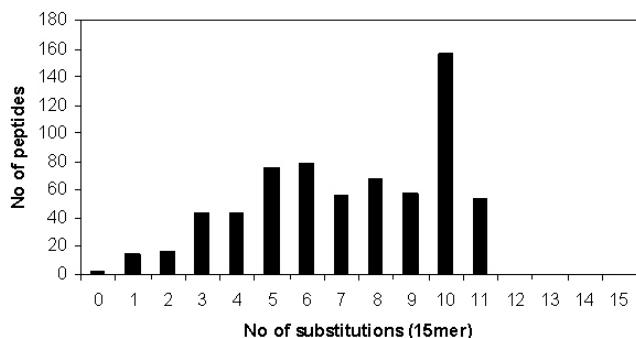
Figure 5
Redundancy profile of predicted Dsg1 peptides to Dsg3 sequences. The number of predicted high-affinity Dsg1 (15 mer) peptide sequences shown as a function of the number of substitutions from the highest identity peptide sequence derived from the Dsg3 proteome derived using a 15 mer sliding window.
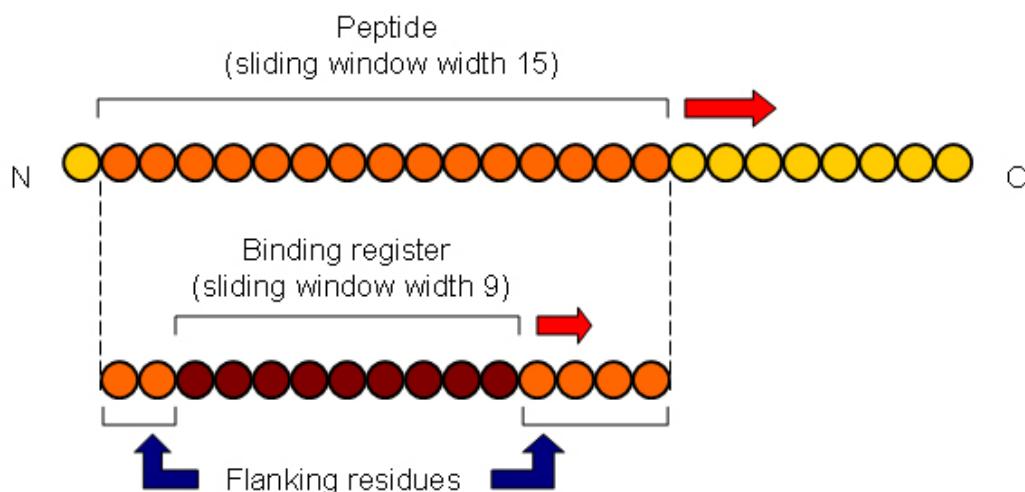
Figure 6
Generation of Dsg1 and Dsg3 peptide sequences for docking into DRB1*0402. Overlapping 15 mer peptides are generated from the Dsg1 and Dsg3 sequences. For each 15 mer peptide, an overlapping sliding window of size nine is applied to generate all combinations of binding registers to be modeled into the binding groove of DRB1*0402.

Methods

The sequence of DRB1*0402 was obtained from IMGT-HLA database [41]. Dsg1 and Dsg3 sequence data were obtained from Swiss-Prot [42]. To identify suitable structural templates in the Protein Data Bank (PDB) [43] for model construction, a sequence similarity search was performed using BLASTP [44] and the highest quality templates (with the best resolution, highest sequence similarity and minimal number of missing residues) were selected among the returned hits. The crystal structure of DRB1*0401 (PDB code 1D5Z) was selected as template for DRB1*0402 (97.9% identity) [33], while the solution structures of C-cadherin (PDB code 1Q5C) and N-cadherin (PDB code 1NCJ) were used for constructing the extracellular domains of both Dsg1 and Dsg3.

Model Building

The program MODELLER [45] was employed for comparative modeling of DRB1*0402, Dsg1 and Dsg3. Each model was constructed by optimally satisfying spatial constraints obtained from the alignment of the template structure with the target sequence and from the CHARMM-22 force field [46], and subsequently relaxed by conjugate gradient minimization, using the program Internal Coordinate Mechanics (ICM) [47,48].

Peptide Docking

Overlapping 15 mer peptides are generated from the Dsg1 and Dsg3 sequences. An overlapping sliding window of size nine is applied to each 15 mer peptide to generate all combinations of binding registers to be modeled into the binding groove of DRB1*0402 (Figure 6). A total of 1035

Dsg1 15 mer peptides (6210 binding registers) and 985 Dsg3 15 mer peptides (5910 binding registers) were generated and used in the current analysis. Docking was performed using a standard protocol as previously described [33-35], consisting of (i) pseudo-Brownian rigid body docking of peptide fragments to the ends of the binding groove, (ii) central loop closure by satisfaction of spatial constraints, (iii) refinement of the backbone and side-chain atoms of the core recognition residues and receptor contact regions within 4.00 Å radius, and (iv) extension of flanking peptide residues by satisfaction of spatial constraints (Figure 7).

Empirical Free Energy Functions

The scoring function presented in this study is based on the free energy potential in ICM [47,48]. Computation of the binding free energy was performed according to previous similar works [36,37] based on the difference between the energy of the solvated complex and the sum of the energy of the solvated receptor and that of the peptide ligand, followed by optimization using experimental IC_{50} values. The DRB1*0402 model optimization, training and testing were described earlier [33].

Immunological Hotspot Prediction

In the present study, 'immunological hotspots' are defined as antigenic regions of up to 30 amino acids based on the sum of predicted binding energies of the top four binders within a window of 30 amino acids [49,50]. Where available, predicted hotspots were validated using available experimentally determined sites (Figure 3).
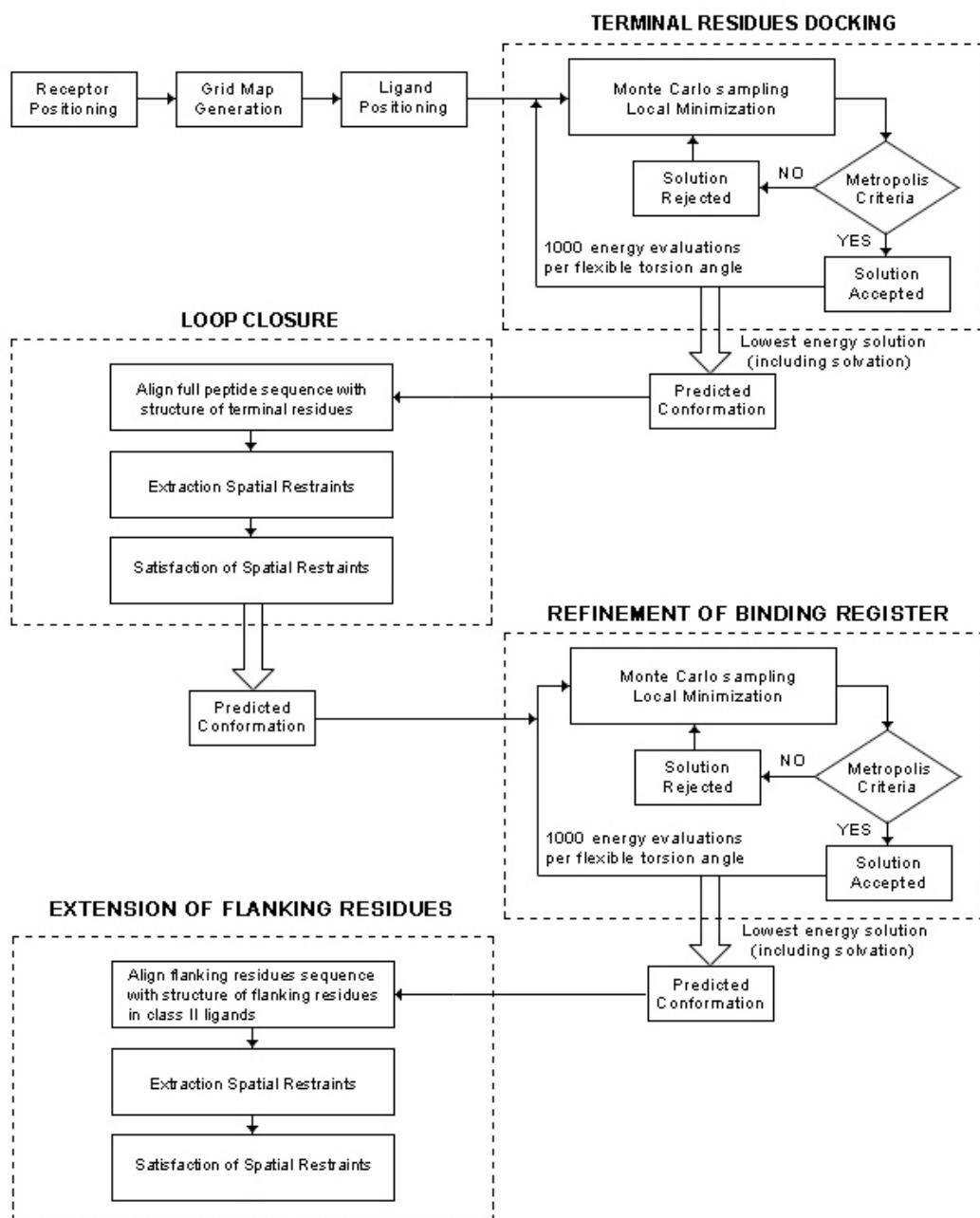
Figure 7
Flowchart of the four-step docking procedure used in this work. Peptides are docked into DRB1*0402 using a four-step procedure consisting of: (i) pseudo-Brownian rigid body docking of peptide fragments to the ends of the binding groove, (ii) central loop closure by satisfaction of spatial constraints, (iii) refinement of the backbone and side-chain atoms of the core recognition residues and receptor contact regions and (iv) extension of flanking peptide residues by satisfaction of spatial constraints.

Accessible Surface Areas

Solvent accessible surface areas were calculated with the program NACCESS [51] by tracing out the maximum permitted van der Waals' contact that is covered by the center of a water molecule (1.40 Å probe radius) as it rolls over the surface of the protein.

Authors' contributions

JCT carried out the computational simulation studies. JCT and AAS participated in the experimental design, data interpretation and drafted the manuscript.

References

1. Beutner EH, Jordon RE: **Demonstration of skin antibodies in sera of pemphigus vulgaris patients by indirect immunofluorescent staining.** *Proc Soc Exp Biol Med* 1964, **117**:505-510.
2. Ishii K, Amagai M, Hall RP, Hashimoto T, Takayanagi A, Gamou S, Shimizu N, Nishikawa T: **Characterization of autoantibodies in pemphigus using antigen-specific enzyme-linked immunosorbent assays with baculovirus-expressed recombinant desmogleins.** *J Immunol* 1997, **159**:2010-2017.
3. Harman KE, Gratian MJ, Shirlaw PJ, Bhogal BS, Challacombe SJ, Black MM: **The transition of pemphigus vulgaris into pemphigus foliaceus: a reflection of changing desmoglein 1 and 3 autoantibody levels in pemphigus vulgaris.** *Br J Dermatol* 2002, **146**:684-687.
4. Miyagawa S, Amagai M, Iida T, Yamamoto Y, Nishikawa T, Shirai T: **Late development of antidesmoglein 1 antibodies in pemphigus vulgaris: correlation with disease progression.** *Br J Dermatol* 1999, **141**:1084-1087.
5. Tchernev G, Orfanos CE: **Antigen mimicry, epitope spreading and the pathogenesis of pemphigus.** *Tissue Antigens* 2006, **68**:280-286.
6. Salato VK, Hacker-Foegen MK, Lazarova Z, Fairley JA, Lin MS: **Role of intracellular epitope spreading in pemphigus vulgaris.** *Clin Immunol* 2005, **116**:54-64.
7. Kalish RS: **Pemphigus vulgaris: the other half of the story.** *J Clin Invest* 2000, **106**:1433-1435.
8. Chan LS: **Epitope spreading in paraneoplastic pemphigus. Autoimmune induction in antibody-mediated blistering skin diseases.** *Arch Dermatol* 2000, **136**:663-664.
9. Vanderlugt CJ, Miller SD: **Epitope spreading.** *Curr Opin Immunol* 1996, **8**:831-836.
10. Chan LS, Vanderlugt CJ, Hashimoto T, Nishikawa T, Zone JJ, Black MM, Wojnarowska F, Stevens SR, Chen M, Fairley JA, Woodley DT, Miller SD, Gordon KB: **Epitope spreading: lessons from autoimmune skin diseases.** *J Invest Dermatol* 1998, **110**:103-109.
11. Lin M-S, Swartz SJ, Lopez A, Ding X, Fernandez-Vina MA, Stastny P, Fairley JA, Diaz LA: **Development and characterization of desmoglein-3 specific T cells from patients with pemphigus vulgaris.** *J Clin Invest* 1997, **99**:31-40.
12. Hertl M, Amagai M, Sundaram H, Stanley J, Ishii K, Katz SI: **Recognition of desmoglein 3 by autoreactive T cells in pemphigus vulgaris patients and normals.** *J Invest Dermatol* 1998, **110**:62-66.
13. Wucherpfennig KW, Yu B, Bhol K, Monos DS, Argyris E, Karr RW, Ahmed AR, Strominger JL: **Structural basis for major histocompatibility complex (MHC)-linked susceptibility to autoimmunity: charged residues of a single MHC binding pocket confer selective presentation of self-peptides in pemphigus vulgaris.** *Proc Natl Acad Sci USA* 1995, **92**:11935-11939.
14. Rizzo C, Fotino M, Zhang Y, Chow S, Spizuoco A, Sinha AA: **Direct characterization of human T cells in pemphigus vulgaris reveals elevated autoantigen-specific Th2 activity in association with active disease.** *Clin Exp Dermatol* 2005, **30**:535-540.
15. Veldman C, Stauber A, Wassmuth R, Uter W, Schuler G, Hertl M: **Dichotomy of autoreactive Th1 and Th2 cell responses to desmoglein 3 in patients with pemphigus vulgaris (PV) and healthy carriers of PV-associated HLA class II alleles.** *J Immunol* 2003, **170**:635-642.
16. Veldman CM, Gebhard KL, Uter W, Wassmuth R, Grotzinger J, Schultz E, Hertl M: **T cell recognition of desmoglein 3 peptides in patients with pemphigus vulgaris and healthy individuals.** *J Immunol* 2004, **172**:3883-3892.
17. Veldman C, Pahl A, Beissert S, Hansen W, Buer J, Dieckmann D, Schuler G, Hertl M: **Inhibition of the transcription factor Foxp3 converts desmoglein 3-specific type 1 regulatory T cells into Th2-like cells.** *J Immunol* 2006, **176**:3215-3222.
18. Hertl M, Eming R, Veldman C: **T cell control in autoimmune bullous skin disorders.** *J Clin Invest* 2006, **116**:1159-1166.
19. Chow S, Rizzo C, Ravitskiy L, Sinha AA: **The role of T cells in cutaneous autoimmune disease.** *Autoimmunity* 2005, **38**:303-317.
20. Lin MS, Swartz SJ, Lopez A, Ding X, Fairley JA, Diaz LA: **T lymphocytes from a subset of patients with pemphigus vulgaris respond to both desmoglein-3 and desmoglein-1.** *J Invest Dermatol* 1997, **109**:734-737.
21. Loiseau P, Lecleach L, Prost C, Lepage V, Busson M, Bastuji-Garin S, Roujeau JC, Charron D: **HLA class II polymorphism contributes to specify desmoglein derived peptides in pemphigus vulgaris and pemphigus foliaceus.** *J Autoimmun* 2000, **15**:67-73.
22. Miyagawa S, Higashimine I, Iida T, Yamashina Y, Fukumoto T, Shirai T: **HLA-DRB1*04 and DRB1*14 alleles are associated with susceptibility to pemphigus among Japanese.** *J Invest Dermatol* 1997, **109**:615-618.
23. Niizeki H, Inoko H, Narimatsu H, Takata H, Sonoda A, Tadakuma T, Ando A, Tsuji K, Hashimoto T, Nishikawa T: **HLA class II antigens are associated with Japanese pemphigus patients.** *Hum Immunol* 1991, **31**:246-50.
24. Ahmed AR, Yunis EJ, Khatri K, Wagner R, Notani G, Awdeh Z, Alper CA: **Major histocompatibility complex haplotype studies in Ashkenazi Jewish patients with pemphigus vulgaris.** *Proc Natl Acad Sci USA* 1990, **87**:7658-7662.
25. Sinha AA, Lopez MT, McDevitt HO: **Autoimmune diseases: the failure of self tolerance.** *Science* 1990, **248**:1380-1388.
26. Sinha AA, Brautbar C, Szafer F, Friedmann A, Tzfoni E, Todd JA, Steinman L, McDevitt HO: **A newly characterized HLA DQ beta allele associated with pemphigus vulgaris.** *Science* 1988, **239**:1026-1029.
27. Ahmed AR, Yunis EJ, Khatri K, Wagner R, Notani G, Awdeh Z, Alper CA: **Major histocompatibility complex haplotype studies in Ashkenazi Jewish patients with pemphigus vulgaris.** *Proc Natl Acad Sci USA* 1990, **87**:7658-7662.
28. Loiseau P, Lecleach L, Prost C, Lepage V, Busson M, Bastuji-Garin S, Roujeau JC, Charron D: **HLA class II polymorphism contributes to specify desmoglein derived peptides in pemphigus vulgaris and pemphigus foliaceus.** *J Autoimmun* 2000, **15**:67-73.
29. Lombardi ML, Mercuro O, Ruocco V, Lo Schiavo A, Lombari V, Guerrera V, Pirozzi G, Manzo C: **Common human leukocyte antigen alleles in Pemphigus Vulgaris and Pemphigus Foliaceus Italian Patients.** *J Invest Dermatol* 1999, **113**:107-110.
30. Gonzalez-Escribano MF, Jimenez G, Walter K, Montes M, Perez-Bernal AM, Rodriguez MR, Conejo-Mir JS, Nunez-Roldan A: **Distribution of HLA class II alleles among Spanish patients with pemphigus vulgaris.** *Tissue Antigens* 1998, **52**:275-278.
31. Glorio R, Rodriguez Costa G, Haas R, Gruber M, Fainboim L, Woscoff A: **HLA Haplotypes and Class II Molecular Alleles in Argentinian Patients with Pemphigus Vulgaris.** *J Cutan Med Surg* 2002, **6**:422-426.
32. Mobini N, Yunis EJ, Alper CA, Yunis JJ, Delgado JC, Yunis DE, Firooz A, Dowlati Y, Bahar K, Gregersen PK, Ahmed AR: **Identical MHC markers in non-Jewish Iranian and Ashkenazi Jewish patients with pemphigus vulgaris: possible common central Asian ancestral origin.** *Hum Immunol* 1997, **57**:62-67.
33. Tong JC, Bramson J, Kanduc D, Sinha AA, Ranganathan S: **Prediction of desmoglein-3 peptides reveals multiple shared T-cell epitopes in HLA DR4- and DR6-associated pemphigus vulgaris.** *BMC Bioinformatics* 2006, **7**:S7.
34. Tong JC, Bramson J, Kanduc D, Chow S, Sinha AA, Ranganathan S: **Modeling the bound conformation of pemphigus vulgaris-associated peptides to MHC class II DR and DQ alleles.** *Immunome Res* 2006, **2**:1.

35. Tong JC, Tan TW, Ranganathan S: **Modeling the structure of bound peptide ligands to major histocompatibility complex.** *Protein Sci* 2004, **13:**2523-2532.
36. Tong JC, Zhang ZH, August JT, Brusic V, Tan TW, Ranganathan S: *In silico* **characterization of immunogenic epitopes presented by HLA-Cw*0401.** *Immunome Res* in press.
37. Tong JC, Zhang GL, Tan TW, August JT, Brusic V, Ranganathan S: **Prediction of HLA-DQ3.2β ligands: Evidence of multiple registers in class II binding peptides.** *Bioinformatics* 2006, **22:**1232-1238.
38. Lucchese A, Mittelman A, Tessitore L, Serpico R, Sinha AA, Kanduc D: **Proteomic definition of a desmoglein linear determinant common to Pemphigus vulgaris and Pemphigus foliaceous.** *J Transl Med* 2006, **4:**37.
39. Lin MS, Swartz SJ, Lopez A, Ding X, Fairley JA, Diaz LA: **T lymphocytes from a subset of patients with pemphigus vulgaris respond to both desmoglein-3 and desmoglein-1.** *J Invest Dermatol* 1997, **109:**734-737.
40. Seishima M, Esaki C, Osada K, Mori S, Hashimoto T, Kitajima Y: **Pemphigus IgG, but not bullous pemphigoid IgG, causes a transient increase in intracellular calcium and inositol 1,4,5-triphosphate in DJM-1 cells, a squamous cell carcinoma line.** *J Invest Dermatol* 1995, **104:**33-37.
41. Robinson J, Waller MJ, Parham P, de Groot N, Bontrop R, Kennedy LJ, Stoehr P, Marsh SG: **IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex.** *Nucleic Acids Res* 2003, **31:**311-314.
42. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S, Schneider M: **The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003.** *Nucleic Acids Res* 2003, **31:**365-370.
43. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissiq H, Shindyalov IN, Bourne PE: **The Protein Data Bank.** *Nucleic Acids Res* 2000, **28:**235-242.
44. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215:**403-410.
45. Sali A, Blundell TL: **Comparative protein modelling by satisfaction of spatial restraints.** *J Mol Biol* 1993, **234:**779-815.
46. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M: **All-atom empirical potential for molecular modeling and dynamics studies of proteins.** *J Phys Chem* 1998, **B102:**3586-3617.
47. Abagyan R, Totrov M, Kuznetsov D: **ICM – a new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation.** *J Comput Chem* 1994, **15:**488-506.
48. Abagyan RA, Totrov M: **Ab initio folding of peptides by the optimal-bias Monte Carlo minimization procedure.** *J Comput Phys* 1999, **151:**402-421.
49. Srinivasan KN, Zhang GL, Khan AM, August JT, Brusic V: **Prediction of class I T-cell epitopes: evidence of presence of immunological hot spots inside antigens.** *Bioinformatics* 2004, **20:**i297-i302.
50. Zhang GL, Khan AM, Srinivasan KN, August JT, Brusic V: **MULTI-PRED: a computational system for prediction of promiscuous HLA binding peptides.** *Nucleic Acids Res* 2005, **33:**W172-W179.
51. Hubbard SJ, Thornton JM: **NACCESS.** Univ Coll, London; 1993.